Some pitfalls in unit root testing

Bakker, B.B.; Koning, Ruud H.

MEMORANDUM

from

INSTITUTE OF ECONOMIC RESEARCH

FACULTY OF ECONOMICS

UNIVERSITY OF GRONINGEN

P.O. Box 800

9700 AV Groningen

Research Memorandum nr. 441

SOME PITFALLS IN UNIT ROOT TESTING

by

Bas B. Bakker
and
Ruud H. Koning *

* The authors thank Theo Dijkstra, Jan Jacobs and Tom Wansbeek for their comments, without implicating them for remaining errors.

September 1991

Abstract

Testing for unit roots is now common practice for economists. The most popular procedure is the approach developed by Dickey and Fuller (1979, 1981), which only requires running appropriately specified regressions. However, application of the Dickey-Fuller procedure requires that the disturbance term is identically independently normally distributed. In this paper we show that the augmented Dickey-Fuller-test (ADF-test) is biased in finite samples if the disturbance term does not follow a white noise process. We will show that for many values of the parameters, the ADF-test accepts the maintained hypothesis of a unit root too often, if the disturbance term follows an AR, MA or ARMA process. Moreover, this problem is aggravated if many lags of Δy_t are included in the testing regression. These points will be made clear both analytically and by simulation evidence.

1 Introduction

In their seminal papers, Dickey and Fuller (1979, 1981) developed a class of test statistics, known as Dickey-Fuller (DF) statistics, to test whether a pure AR(1) process has a unit root. Consider the following data generating process:

$$y_t = \rho y_{t-1} + \varepsilon_t, \tag{1}$$

with $\varepsilon_t \sim I\mathcal{N}(0, \sigma_\varepsilon^2)$ distributed. Dickey and Fuller suggest to rewrite equation (1) as

$$\Delta y_t = \gamma y_{t-1} + \varepsilon_t, \tag{2}$$

and to use the t–statistic of $\gamma \equiv \rho - 1$ to test whether γ equals 0, and hence, ρ equals 1. Note that, if $\rho < 1$, γ will be negative. However, under the null hypothesis that $\rho = 1$, this statistic does not follow the standard Student t–distribution. Dickey and Fuller derive the limit distributions and present empirical distributions, based on Monte Carlo simulations.

If the data generating process is AR(p) instead of AR(1), *i.e.*:

$$y_t = \rho_1 y_{t-1} + \rho_2 y_{t-2} + \cdots + \rho_p y_{t-p} + \varepsilon_t, \tag{3}$$

with $\varepsilon_t \sim I\mathcal{N}(0, \sigma_\varepsilon^2)$, Dickey and Fuller (1981) suggest rewriting equation (3) as

$$\Delta y_t = \gamma y_{t-1} + \theta_1 \Delta y_{t-1} + \theta_2 \Delta y_{t-2} + \cdots + \theta_{p-1} \Delta y_{t-p-1} + \varepsilon_t, \tag{4}$$

where $\gamma = \sum_{i=1}^{p} \rho_i - 1$ and $\theta_j = -\sum_{i=j+1}^{p} \rho_i$ and to test whether γ equals 0, by means of the t–statistic of γ. This test is known as the Augmented Dickey Fuller test which will be referred to as the ADF-test in the sequel.. The critical values are the same as those of the DF statistics.

In the derivation of their limiting distributions, Dickey and Fuller assume that ε_t is uncorrelated. Said and Dickey (1984) consider a more general process than the process given in equation (1):

$$\begin{aligned} y_t &= \rho y_{t-1} + \varepsilon_t \\ \varepsilon_t &= \alpha_1 \varepsilon_{t-1} + \alpha_2 \varepsilon_{t-2} + \cdots + \alpha_k \varepsilon_{t-k} + \eta_t + \beta_1 \eta_{t-1} + \cdots + \beta_q \eta_{t-q} \end{aligned}$$

and show that the ADF-test can be used if enough lags are included.[1] Their results, however, are based on asymptotic arguments. Simulation evidence in Molinas (1986) and Schwert (1987, 1989) shows that this test is affected by the process generating the data in finite samples. In particular, if ε_t follows a MA(1) process, *i.e.*:

$$\varepsilon_t = \eta_t + \beta \eta_{t-1}$$

with β close to one, the ADF statistics have critical values that are far below the ones tabulated in the Dickey-Fuller papers, which too frequently implies that these tests will lead to the conclusion that economic data are stationary.

Another approach to testing for unit roots, which allows for a more general stochastic process of ε_t, has been developed by Phillips (1987) and Phillips and Perron (1988). However, computation of their statistics is not as straightforward as computation of the

[1] To be more precise, if p lags are included with $p = \mathcal{O}(T^{1/3})$, with T the number of observations.

Dickey-Fuller statistics, which is probably the reason why the approach by Dickey and Fuller has been increasingly popular with applied economists.

In this paper, we will show that the ADF statistic often leads to the wrong conclusion if ε_t is correlated. There is a tendency to accept the null-hypothesis of a unit-root in the data-generating process too often if ε_t is correlated and because the test has a rather low power against alternatives which lie close to the null-hypothesis. This point will be made clear analytically as well by simulation evidence.

The paper is organized as follows. In section 2 it is shown why the ADF test is biased, if ε_t is correlated. In section 3, we present the size and the power of the ADF-test under various specifications of ε_t, where we include four lags of Δy_t. In section 4 we look at the same quantities, but there we use only one lag of Δy_t in the testing regression. Section 5 concludes.

2 Why ADF-Tests Are Biased

Consider the process

$$\begin{aligned} y_t &= \rho_1 y_{t-1} + \rho_2 y_{t-2} + \cdots + \rho_p y_{t-p} + \varepsilon_t \\ \varepsilon_t &= \alpha_1 \varepsilon_{t-1} + \alpha_2 \varepsilon_{t-2} + \cdots + \alpha_k \varepsilon_{t-k} + \eta_t + \beta_1 \eta_{t-1} + \cdots + \beta_q \eta_{t-q}. \end{aligned}$$

We will call this process an ARMA$(k,0) \times (p,q)$ process. The standard Dickey-Fuller test tests whether ρ_1 equals 1 in an ARMA$(0,0)\times(1,0)$ process, which is just the process given in equation (1). The ADF-tests test whether ρ_1 is equal to 1 in all other ARMA-$(k,0) \times (p,q)$ specifications. In this section, we will elaborate on three such processes, viz. ARMA$(1,0) \times (1,0)$, ARMA$(0,0) \times (1,1)$ and ARMA$(1,0) \times (1,1)$.

ARMA$(1,0) \times (1,0)$

Consider the model

$$y_t = \rho y_{t-1} + \varepsilon_t \tag{5}$$

$$\varepsilon_t = \alpha \varepsilon_{t-1} + \eta_t. \tag{6}$$

Rewrite equation (6) as

$$\varepsilon_t = \frac{\eta_t}{1 - \alpha L}, \tag{7}$$

where L is the lag operator, *i.e.*, $Lx_t \equiv x_{t-1}$. Substitute equation (7) in equation (5):

$$(1 - \alpha L)(1 - \rho L) y_t = \eta_t,$$

which can be rewritten as

$$y_t = (\alpha + \rho) y_{t-1} - \alpha \rho y_{t-2} + \eta_t. \tag{8}$$

Subtracting y_{t-1} on both sides of equation (8) yields after some manipulating:

$$\Delta y_t = (\rho - 1)(1 - \alpha) y_{t-1} + \alpha \rho \Delta y_{t-1} + \eta_t. \tag{9}$$

We now see that the coefficient of y_{t-1} in equation (9) does not equal $(\rho-1)$ but $(\rho-1)(1-\alpha)$. Depending on the sign of α, the null hypothesis that ρ equals 1 will be rejected too often or too seldom. In economic practice, it is often seen that $\alpha > 0$, and hence, if one interprets the estimate of the coefficient of Δy_{t-1} as an estimate for $\rho-1$, an ADF-test on the coefficient of y_{t-1} in equation (9) will too often lead to acceptance of the null hypothesis that $\rho = 1$, as the test is one-sided.

ARMA$(0,0) \times (1,1)$

Next we consider the data generating process

$$
\begin{aligned}
y_t &= \rho y_{t-1} + \varepsilon_t &(10)\\
\varepsilon_t &= \eta_t - \beta\eta_{t-1} &(11)
\end{aligned}
$$

If we substitute equation (11) in equation (10) we obtain:

$$\frac{1-\rho L}{1-\beta L} y_t = \eta_t,$$

or

$$(1-\rho L)\left(\sum_{\tau=0}^{\infty} \beta^\tau L^\tau\right) y_t = \eta_t,$$

which can be rewritten as

$$\left(1 + (\beta-\rho)L + \beta(\beta-\rho)L^2 + \beta^2(\beta-\rho)L^3 + \sum_{\tau=4}^{\infty}(\beta-\rho)\beta^{\tau-1}L^\tau\right) y_t = \eta_t. \qquad (12)$$

Manipulating equation (12) yields:

$$
\begin{aligned}
\Delta y_t = {} & (\rho-1)(1+\beta)y_{t-1} - \rho\beta\Delta y_{t-1} - \beta^2\Delta y_{t-2}\\
& + (\beta^2(\rho-1)(1+\beta) - \rho\beta^3)\Delta y_{t-3}\\
& + \sum_{i=2}^{\infty}(\lambda_{2i}\Delta y_{t-2i} + \lambda_{2i+1}\Delta y_{t-2i-1}) + \eta_t, \qquad (13)
\end{aligned}
$$

where

$$
\begin{aligned}
\lambda_{2i} &= (\rho-1)(1+\beta)(\beta^2 + \cdots + \beta^{2i-2}) - \beta^{2i}\\
\lambda_{2i+1} &= (\rho-1)(1+\beta)(\beta^2 + \cdots + \beta^{2i}) - \rho\beta^{2i+1}.
\end{aligned}
$$

The coefficient of y_{t-1} is now equal to $(\rho-1)(1+\beta)$. We now see why the null hypothesis is rejected too often if $\beta < 0$, in accordance with the findings of Schwert (1989).

ARMA$(1,0) \times (1,1)$

As a generalization of the models described above, consider an ARMA$(1,0) \times (1,1)$ model:

$$
\begin{aligned}
y_t &= \rho y_{t-1} + \varepsilon_t &(14)\\
\varepsilon_t &= \alpha\varepsilon_{t-1} + \eta_t - \beta\eta_{t-1}. &(15)
\end{aligned}
$$

We can rewrite these equations as

$$(1-\alpha L)(1-\rho L)y_t = (1-\beta L)\eta_t,$$

or as

$$\frac{(1-\alpha L)(1-\rho L)}{1-\beta L} y_t = \eta_t.$$

Expanding the denominator in a power series, we obtain:

$$(1-(\alpha+\rho)L+\alpha\rho L^2)\left(\sum_{\tau=0}^{\infty}\beta^\tau L^\tau\right) y_t = \eta_t.$$

Tedious manipulations finally yield:

$$\begin{aligned}\Delta y_t = {} & (\rho-1)(1-\alpha)(1+\beta)y_{t-1} + (\alpha\rho-(\alpha+\rho)\beta+\alpha\rho\beta)\Delta y_{t-1} \\ & -(\beta^2-\alpha\rho\beta)\Delta y_{t-2} \\ & +((\rho-1)(1-\alpha)(1+\beta)\beta^2-(\alpha+\rho)\beta^3+\alpha\rho\beta^2(1+\beta))\Delta y_{t-3} \\ & +\sum_{i=2}^{\infty}(\lambda_{2i}\Delta y_{t-2i}+\lambda_{2i+1}\Delta y_{t-2i-1})+\eta_t, \end{aligned} \tag{16}$$

with

$$\begin{aligned}\lambda_{2i} &= (\rho-1)(1-\alpha)(1+\beta)(\beta^2+\cdots+\beta^{2i-2})-\beta^{2i}+\alpha\rho\beta^{2i-1}, \\ \lambda_{2i+1} &= (\rho-1)(1-\alpha)(1+\beta)(\beta^2+\cdots+\beta^{2i})+\alpha\rho\beta^{2i}(\beta+1) \\ & \quad -(\alpha+\rho)\beta^{2i+1}.\end{aligned}$$

Again, the coefficient of y_{t-1} can not be interpreted as an estimate of $\rho-1$. Moreover, note that equation (9) can be obtained from equation (16) by putting $\beta=0$, and that equation (13) can be obtained by putting $\alpha=0$. Equation(16) suggests that the ADF-test will give reasonable results if $(1-\alpha)(1+\beta)\approx 1$. In that case, the estimate of the coefficient of y_{t-1} will be a reasonable estimate for $\rho-1$.

3 The size and the power of the augmented Dickey-Fuller test, four lags

In order to get some idea of the size and the power of unit root tests for various specifications of the disturbance term, we have performed some simulations. It is well-known that the nominal size of a unit root test changes if the data generating process is, say, an ARMA$(0,0)\times(1,1)$ process instead of a pure AR(1) process. However, there appears to be little Monte-Carlo evidence on the performance of the ADF-test if the true data generating process is a ARMA$(1,0)\times(1,0)$ process. In fact, we have simulated a more general model:

$$\begin{aligned} y_t &= \rho y_{t-1}+\varepsilon_t \\ \varepsilon_t &= \alpha\varepsilon_{t-1}+\eta_t+\beta\eta_{t-1} \end{aligned} \tag{17}$$

For various values of ρ, α and β, we have tested the hypothesis that $\rho = 1$, using the Augmented Dickey-Fuller test, based on the regression equation

$$\Delta y_t = \gamma y_{t-1} + \delta_1 \Delta y_{t-1} + \delta_2 \Delta y_{t-2} + \delta_3 \Delta y_{t-3} + \delta_4 \Delta y_{t-4} + \zeta_t, \tag{18}$$

(*cf.* equation (9) and equation (13)). The presence of a unit root was tested for using the t-value of $\hat{\gamma}$, which follows a $\hat{\tau}$ distribution under the null hypothesis $\rho = 1$ (see Fuller (1976), table 8.5.2). We tested at a significance level of 5%.

It is clear that too many lags are included in equation (18) if the data generating process is an ARMA(0,0) $\times$ (1,0) process or an ARMA(1,0) $\times$ (1,0) process (these cases correspond to $\alpha = \beta = 0$ and $\beta = 0$ in equation (17) respectively). This may lead to a bias in the estimated standard error of $\hat{\gamma}$. It is well known that the standard error of an estimator in an OLS-model is inflated if too many irrelevant regressors are included. This implies that the t-value is biased towards 0 and that the null-hypothesis will be rejected too seldom.

The simulation results are presented in the tables I-1 to I-5 below. They are based on 1000 replications of a sample of 100 observations. In fact, each time a sample of 400 observations was generated, but only the last 100 observations were used in the analysis. This was to avoid initial value problems. As starting values of the process, we have taken $y_0 = \varepsilon_0 = \eta_0 = 0$, and the η_t's are independently standard normally distributed.

	$\alpha =$						
	−0.75	−0.50	−0.25	0	0.25	0.50	0.75
$\beta =$							
−0.75	0.000	0.000	0.000	0.002	0.002	0.007	0.081
−0.50	0.006	0.009	0.020	0.029	0.050	0.095	0.249
−0.25	0.046	0.046	0.059	0.082	0.082	0.130	0.281
0	0.062	0.072	0.071	0.094	0.114	0.133	0.280
0.25	0.074	0.075	0.084	0.098	0.110	0.175	0.311
0.50	0.069	0.110	0.099	0.126	0.148	0.193	0.360
0.75	0.112	0.123	0.134	0.153	0.223	0.257	0.425

Table I-1: Probability of accepting H_0, $\rho = 0.80$

	$\alpha =$						
	−0.75	−0.50	−0.25	0	0.25	0.50	0.75
$\beta =$							
−0.75	0.000	0.000	0.002	0.000	0.010	0.024	0.198
−0.50	0.038	0.075	0.064	0.094	0.126	0.181	0.371
−0.25	0.114	0.136	0.138	0.166	0.183	0.233	0.377
0	0.155	0.160	0.175	0.192	0.197	0.273	0.410
0.25	0.180	0.193	0.184	0.216	0.230	0.286	0.381
0.50	0.177	0.196	0.219	0.213	0.254	0.320	0.472
0.75	0.169	0.211	0.253	0.306	0.342	0.390	0.501

Table I-2: Probability of accepting H_0, $\rho = 0.85$

	$\alpha =$						
	−0.75	−0.50	−0.25	0	0.25	0.50	0.75
$\beta =$							
−0.75	0.011	0.013	0.021	0.040	0.079	0.138	0.409
−0.50	0.213	0.250	0.283	0.268	0.357	0.407	0.528
−0.25	0.367	0.373	0.374	0.379	0.394	0.464	0.553
0	0.372	0.369	0.378	0.394	0.410	0.461	0.529
0.25	0.403	0.376	0.395	0.425	0.416	0.471	0.543
0.50	0.395	0.420	0.428	0.406	0.460	0.508	0.587
0.75	0.442	0.447	0.470	0.514	0.556	0.550	0.658

Table I-3: Probability of accepting H_0, $\rho = 0.90$

	$\alpha =$						
	−0.75	−0.50	−0.25	0	0.25	0.50	0.75
$\beta =$							
−0.75	0.189	0.192	0.229	0.306	0.377	0.469	0.704
−0.50	0.607	0.596	0.653	0.636	0.677	0.679	0.766
−0.25	0.664	0.691	0.713	0.682	0.690	0.704	0.743
0	0.726	0.686	0.696	0.712	0.706	0.699	0.745
0.25	0.678	0.680	0.706	0.704	0.699	0.719	0.748
0.50	0.676	0.712	0.723	0.735	0.746	0.735	0.759
0.75	0.707	0.745	0.736	0.782	0.821	0.796	0.793

Table I-4: Probability of accepting H_0, $\rho = 0.95$

	$\alpha =$						
	−0.75	−0.50	−0.25	0	0.25	0.50	0.75
$\beta =$							
−0.75	0.867	0.898	0.906	0.907	0.924	0.948	0.948
−0.50	0.953	0.942	0.945	0.941	0.946	0.954	0.954
−0.25	0.946	0.961	0.948	0.967	0.953	0.954	0.946
0	0.960	0.941	0.950	0.952	0.949	0.950	0.953
0.25	0.968	0.946	0.956	0.956	0.954	0.945	0.951
0.50	0.957	0.943	0.940	0.958	0.954	0.957	0.947
0.75	0.962	0.946	0.951	0.958	0.939	0.946	0.952

Table I-5: Probability of accepting H_0, $\rho = 1$

From the simulations, we see that the null hypothesis of a unit root is accepted too often if the parameters α and/or β of the ARMA process generating the disturbances are large. For example, if $\rho = 0.95$, $\alpha = 0$ and $\beta = 0.50$, we find that the test statistic exceeds the critical value with probability 0.735, even though the data generating process does not have a unit root.

Finally, we see that the true significance levels are not extremely sensitive for misspecification of the process generating the disturbances, unless the MA-parameter β is large and negative. This corresponds to the findings of Schwert (1989).

4 The size and power of the augmented Dickey-Fuller test, one lag

Instead of estimating equation (18), one can also estimate this equation with one lag Δy_{t-1}, which is the appropriate specification of the disturbances follow a pure AR(1) process (see equation (9)). The probabilities of accepting the maintained hypothesis of a unit root process in y_t are given in table II-1 to table II-5.

	$\alpha =$						
	−0.75	−0.50	−0.25	0	0.25	0.50	0.75
$\beta =$							
−0.75	0.000	0.000	0.000	0.000	0.000	0.000	0.012
−0.50	0.000	0.000	0.000	0.000	0.000	0.011	0.273
−0.25	0.000	0.000	0.000	0.003	0.008	0.064	0.277
0	0.000	0.000	0.008	0.009	0.017	0.053	0.179
0.25	0.023	0.021	0.014	0.010	0.014	0.016	0.084
0.50	0.026	0.007	0.005	0.003	0.002	0.011	0.041
0.75	0.013	0.007	0.004	0.004	0.007	0.004	0.023

Table II-1: Probability of accepting H_0, $\rho = 0.80$

	$\alpha =$						
	−0.75	−0.50	−0.25	0	0.25	0.50	0.75
$\beta =$							
−0.75	0.000	0.000	0.000	0.000	0.000	0.000	0.000
−0.50	0.000	0.000	0.000	0.002	0.004	0.069	0.448
−0.25	0.000	0.002	0.011	0.023	0.067	0.185	0.461
0	0.049	0.046	0.051	0.057	0.076	0.140	0.277
0.25	0.128	0.082	0.062	0.053	0.052	0.076	0.122
0.50	0.141	0.071	0.034	0.029	0.024	0.036	0.085
0.75	0.055	0.025	0.026	0.021	0.019	0.021	0.056

Table II-2: Probability of accepting H_0, $\rho = 0.85$

We see that the power of the test is increased significantly. To return to our example (which is by no means an exception), we see that the probability of accepting the (false) null hypothesis of a unit root drops from 0.735 to 0.530. In general, there is less tendency

	$\alpha =$ -0.75	-0.50	-0.25	0	0.25	0.50	0.75
$\beta =$							
−0.75	0.000	0.000	0.000	0.000	0.000	0.007	0.290
−0.50	0.000	0.000	0.003	0.022	0.086	0.269	0.713
−0.25	0.037	0.068	0.115	0.171	0.254	0.402	0.644
0	0.258	0.268	0.271	0.290	0.302	0.322	0.465
0.25	0.418	0.324	0.282	0.248	0.205	0.192	0.288
0.50	0.401	0.297	0.211	0.164	0.136	0.129	0.193
0.75	0.286	0.181	0.143	0.113	0.093	0.104	0.155

Table II-3: Probability of accepting H_0, $\rho = 0.90$

	$\alpha =$ -0.75	-0.50	-0.25	0	0.25	0.50	0.75
$\beta =$							
−0.75	0.000	0.000	0.000	0.001	0.015	0.148	0.690
−0.50	0.033	0.081	0.154	0.241	0.465	0.671	0.888
−0.25	0.364	0.414	0.486	0.597	0.678	0.769	0.843
0	0.650	0.668	0.673	0.677	0.683	0.682	0.721
0.25	0.785	0.727	0.673	0.639	0.587	0.582	0.565
0.50	0.754	0.672	0.599	0.530	0.472	0.440	0.464
0.75	0.686	0.581	0.529	0.477	0.430	0.387	0.408

Table II-4: Probability of accepting H_0, $\rho = 0.95$

	$\alpha =$ -0.75	-0.50	-0.25	0	0.25	0.50	0.75
$\beta =$							
−0.75	0.527	0.566	0.634	0.817	0.798	0.898	0.949
−0.50	0.813	0.863	0.893	0.909	0.943	0.964	0.941
−0.25	0.933	0.945	0.937	0.953	0.943	0.952	0.951
0	0.955	0.949	0.962	0.959	0.941	0.945	0.954
0.25	0.954	0.953	0.959	0.951	0.947	0.948	0.928
0.50	0.959	0.955	0.946	0.938	0.943	0.934	0.932
0.75	0.957	0.950	0.943	0.938	0.931	0.916	0.917

Table II-5: Probability of accepting H_0, $\rho = 1$

to accept the null hypothesis when one lag is included in the testing regression than if four lags are included. These findings cast doubts on the recommendations of Schwert to include many lags, in order to ensure that the residuals follow a white noise process.

5 Conclusion

In this paper, we have looked at the consequences of misspecification of the stochastic process of the disturbances in testing for unit roots. We have seen that the ADF-test is biased towards accepting the null hypothesis of a unit root, if the disturbances do not follow a white noise process. Inclusion of many lagged differences in the testing regression may ensure that disturbances do follow a white noise process, but this strategy will also lower the power of the Dickey-Fuller test.

References

Dickey, D.A. and W.A. Fuller (1979), "Distribution of estimators for autoregressive time series with a unit root", *Journal of the American Statistical Association*, **74**, 427–431.

Dickey, D.A. and W.A. Fuller (1981), "Likelihood ratio statistics for autoregressive time series with a unit root", *Econometrica*, **49**, 1057–1072.

Fuller, W.A. (1976), *Introduction to Statistical Time Series*, John Wiley & Sons, New York.

Molinas, C. (1986), "A note on spurious regression with integrated moving average errors", *Oxford Bulletin of Economics and Statistics*, **48**, 279–282.

Phillips, P.C.B. (1987), "Time series regression with a unit root", *Econometrica*, **55**, 277–302.

Phillips, P.C.B. and P. Perron (1988), "Testing for a unit root in time series regression", *Biometrika*, **75**, 335–346.

Said, S.E. and D.A. Dickey (1984), "Testing for unit roots in autoregressive-moving average models of unknown order", *Biometrika*, **71**, 599–607.

Schwert, G.W. (1987), "The effects of model specification on tests for unit roots in macroeconomic data", *Journal of Monetary Economics*, **20**, 73–103.

Schwert, G.W. (1989), "Tests for unit roots: A monte carlo investigation", *Journal of Business & Economic Statistics*, **7**, 147–159.

A Appendix

All simulation were performed in the matrix language GAUSS (version 2.01) on a Hewlett-Packard 286/12 personal computer. The procedures used to calculate the entries in tables I-1 tp I-5 and tables II-1 t0 II-5 is listed below.

```
nrep=1000;
nobs=100;
output file=c:\gauss\results\simuleer.bak reset;
print "I-3";
print "time string";
print timestr(0);
print "aantal replicaties";
print nrep;
print "aantal observaties";
print nobs;
rho=0.90;
theta=-0.75;
   do until (theta == 0);
   gam=-0.75;
   do until (gam == 1);
      {h0,h1}=augm4(nrep,nobs,rho,gam,theta);
      print "vier lags";
      print "rho";
      print rho;
      print "gamma";
      print gam;
      print "theta";
      print theta;
      print "perc. H0";
      print h0;
      print "perc H1";
      print h1;
      print " ";
      print timestr(0);
      gam=gam+0.25;
  endo;
  theta=theta+0.25;
endo;
closeall;
end;
```

```
proc 2=augm4(nrep,nobs,rho,gam,theta);
  local datafile,dt,data,ytmin1,yt,b,e,varb,nobs,t,rho,gam,theta,
  dy,i,nrep,h0,h1,x,T,invxx,dytmin2,dytmin3,dytmin4,dytmin1;
  h0=0;
  h1=0;
  i=1;
  do until (i==(nrep+1));
    data=tijdreek(nobs,rho,gam,theta);
    yt=data[6:100,1];
    ytmin1=data[5:99,1];
    dytmin1=ytmin1-data[4:98,1];
    dytmin2=data[4:98,1]-data[3:97,1];
    dytmin3=data[3:97,1]-data[2:96,1];
    dytmin4=data[2:96,1]-data[1:95,1];
    T=rows(ytmin1);
    x=ytmin1~dytmin1~dytmin2~dytmin3~dytmin4;
    dy=yt-ytmin1;
    invxx=inv(x'x);
    b=invxx*x'*dy;
    e=dy-x*b;
    varb=(e'e/(T-5))*invxx;
    t=b[1,1]/sqrt(varb[1,1]);
    h0=h0+(t>-1.95);
    h1=h1+(t<=-1.95);
    i=i+1;
  endo;
retp(h0/nrep, h1/nrep);
endp;
```

```
proc 1=tijdreek(nobs,rho,gam,theta);
  local file,filenaam,k,y,teller,aantal,ytmin1,ztmin1,etmin1,yt,zt,
        et,t,seed;
  seed=round(nobs*rndu(1,1));
  teller=0;
  aantal=100;
  do until (teller.>=nobs);
    ytmin1=0; /*rndn(1,1);*/
    etmin1=0; /*rndn(1,1);*/
    ztmin1=0; /*rndn(1,1);*/
    y=zeros(3*aantal,1);
    t=1;
    do until (t==(3*aantal+1));
       et=rndn(1,1);
```

```
      zt=gam*ztmin1+et+theta*etmin1;
      y[t,1]=rho*ytmin1+zt;
      ytmin1=y[t,1];
      ztmin1=zt;
      etmin1=et;
      t=t+1;
    endo;
    y=y[200:300,1];
    teller=teller+aantal;
  endo;
  retp(y);
endp;
```

Papers issued in these series from 1991 onwards: [1])

404. Herman J. de Jong, Infrastructure and inland transport in the Netherlands during the first railway years, 1800-1880 (1991).
405. Hans van Ees, Does money affect money prices? (1991).
406. Bas Bakker, Het Netto Buitenlands Actief van de banken in de periode 84-90 (1991).
407.* Claude Wetta, Marché et mobilité de l'emploi au Burkina Faso (1991).
408. Dirk Pilat, Levels of real output and labour productivity by industry of origin, a comparison of Japan and the United States, 1975 and 1970-1987 (1991).
409. Piet Boomsma, Jan Oosterhaven and Anne van der Veen, Construction of a representative Dutch bi-regional input-output table (1991).
410. H.J. Brinkman, J.W. Drukker, S.J. Stuurop, De sociale afkomst der Nederlandse lotelingen in de 19e eeuw (1991).
411. Elmer Sterken, Models of the business cycle; a survey (1991).
412. Bas Bakker, Is debt neutral when firms are cash-in-advance constrained? (1991).
413. Arend Buitendam, Decentralization and the governance of employment relationships: human resource management between labour market and organization (1991).
414. Ruud H. Koning and Geert Ridder, Discrete choice and stochastic utility maximization (1991).
415. L. Schoonbeek, The instantaneous speed of adjustment assumption and dominant diagonals (1991).
416. Eijte W. Foekens and Peter S.H. Leeflang, Comparing scanner data with traditional store audit data (1991).
417. S. Brakman, J. de Haan, C.J. Jepma, De reële effectieve wisselkoers van de gulden (1991).
418. Cars H. Hommes, Periodic, quasi-periodic and chaotic dynamics in a simple macro model with Hicksian nonlinearities (extended version) (1991).
419. Gerard Sierksma, Valeriu Soltan, Tudor Zamfirescu, Invariance of convex sets under linear transformations (1991).
420. Hans ten Kate, Tom van der Hoeven and Gerard Sierksma, Transporting and mixing gasses with different qualities (1991).
421. Gerard J. van den Berg, Explicit expressions for the reservation wage path and the unemployment duration density in nonstationary job search models (1991).
422. C. Praagman, J.M. Soethoudt, A note on finite memory in locally square integrable systems (1991).
423. J.S. Cramer, G. Ridder, Testing for equivalent alternatives in discrete choice models (1991).
424. Vladimir G. Deineko and René van Dal, Solvable cases of the euclidean traveling salesman problem (1991).
427. Gerard J. van den Berg, Maarten Lindeboom, Geert Ridder, Attrition in longitudinal panel data, and the empirical analysis of dynamic labour market behaviour (1991).
428. Günter Knieps, Europe's air traffic control crisis analysis and proposals for reform (1991).

---

[1]) These research memoranda can be ordered at the Institute of Economic Research, University of Groningen, P.O. Box 800, 9700 AV Groningen.

Memoranda marked * are out of print!

429. Jack A.A. van der Veen, René van Dal and Gerard Sierksma, The symmetric circulant traveling salesman problem (1991).
430. Maarten Lindeboom, Gerard J. van den Berg, Heterogeneity in models for bivariate survival: the importance of the mixing distribution (1991).
431. J.W. Nieuwenhuis, A Systems-theoretic approach to rational expectations models (1991).
432. Gerard J. van den Berg, Wage dispersion and mobility (1991).
433. Henk W. Plasmeijer, Een exercitie in dismal science. Ricardiaanse thema's in de milieu-economie (1991).
434. W. Voorhoeve en H.W.A. Dietzenbacher, Het model GRECON 91-D. Voorspellingen voor 1991 en 1992 (1991).
435. Dirk Pilat, Productivity growth in South Korean manufacturing - a comparative perspective, 1953-1988 (1991).
436. Jan Roelf Bult, Peter S.H. Leeflang, Dick R. Wittink, Testing causality between marketing variables using scanner data (1991).
437. Bas Bakker, Do taxes stimulate saving? (1991).
438. Peter A.G. van Bergeijk and Robert Lensink, Trade, capital, and the transition in Central Europe (1991).
439. Charles B. Blankart and Günter Knieps, Path dependence, network externalities and standardization (1991).
440. Dirk Pilat and D.S. Prasada Rao, A multilateral approach to international comparisons of real output, productivity and purchasing power parities in manufacturing (1991).
441. Bas B. Bakker and Ruud H. Koning, Some pitfalls in unit root testing (1991).